

DEC 01 2006

Serial No. 10/729,417  
Wakefield, et al.  
Filed December 5, 2003Response to Office Action  
Examiner: Phuong Thao Cao  
Group Art Unit: 2164**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

**INVENTORS:** Todd D. Wakefield and David L. Bean  
**TITLE:** Methods and Products for Integrating Mixed Format Data  
**FILING DATE:** December 5, 2003 **EXAMINER NAME:** Phuong Thao Cao  
**SERIAL NO.:** 10/729,417 **GROUP ART UNIT:** 2164  
**DOCKET:** 5130 P

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450**REMARKS**

A terminal disclaimer is attached to overcome the double patenting rejection.  
The claim objection has been addressed through the amendment above.  
The section 101 rejection has been addressed through the amendment above. The claims relate to processing data on real world computer-readable storage media.  
The section 112 rejection has been addressed through the amendment above.

Applicant traverses the section 102 and 103 rejections.

On page 10 of the office action, the examiner says the following:

"Haug, et al. teach: whereinsaid instructions are further executable to perform the function of applying caseframes while performing said interpreting the free text" (see [column 5, lines 40-65] wherein a syntactic parsetree as disclosed is equivalent to Applicant's "caseframes").

In short, the examiner equates caseframes with a parsetree. That is incorrect. A caseframe is

Serial No. 10/729,417  
Wakefield, et al.  
Filed December 5, 2003

**Response to Office Action**

Examiner: Phuong Thao Cao  
Group Art Unit: 2164

something can be applied to a parse tree – not a parse tree itself. Paragraph 12 of Applicant's patent application explains caseframes:

"Caseframes, generally speaking, are patterns that identify a particular linguistic construction and an element of that construction to be extracted. A syntactic caseframe, for example, may be applied to a parsed sentence to identify a clause that contains a subject and an active voice verb, and to extract the subject noun phrase. A syntactic caseframe often also uses lexical filters to constrain its identification process. For example, a user might want to extract the names of litigation plaintiffs in legal documents by creating a caseframe that extracts the subjects of a single active voice verb, sue. Other caseframe types may be fashioned, such as thematic role caseframes that apply their patterns, not to syntactic constructions, but thematic role relationships. More than one caseframe may apply to a sentence. If desired, a selection process may be utilized to reduce the number of caseframes that apply to a particular sentence, although under many circumstances that will not be desirable nor necessary."

Haug does not teach this concept. Therefore Haug does not teach caseframes or their use so Haug does not anticipate the claims reciting use of caseframes. Thus all claims that recite use of caseframes are novel and nonobvious.

On page 9 of the office action, the Examiner takes the position that Haug teaches "interpreting the free text of the unstructured data to produce a set of construed data reflecting at least one relational fact conveyed in the free text, each construed datum relatable to a data tuple of the structured data." The Examiner points out Haug's use of ICD9 codes. However, Applicant has clarified its claims through the amendment above that Applicant's invention uses multi-dimensional relational facts. An ICD9 code is a one dimensional relational fact – by definition it is a code. However, Applicant's multi-dimensional relational facts include multiple attributes. For example, consider the following sentence:

"The bolt on the underside of the transmission case was cracked due to heat."

A concept tag applied to this by the system of the Haug patent would indicate a single concept – such as "transmission failure," or at best "heat-related transmission failure." It would not provide multi-dimensional attributes regarding the sentence.

Serial No. 10/729,417  
Wakefield, et al.  
Filed December 5, 2003

**Response to Office Action**

Examiner: Phuong Thao Cao  
Group Art Unit: 2164

The present invention, however, uses relational fact extraction that encompasses multiple dimensions – i.e., what failed = bolt; how did it fail = cracking; where it failed = underside of transmission case; why it failed = heat. In other words, the claimed invention uses relational fact extractions that include multiple potential dimensions or attributes going to who, what, when, where, why and how. This facilitates multi-dimensional analysis. In contrast, a concept tag (Haug) is one dimensional, and cannot support multi-dimensional analysis as envisioned by the invention. This is an extremely important distinction because without a way to handle multi-dimensional attributes, Haug loses most of the meaning in the natural language free text, while the present invention preserves that meaning for use in analytic processing. This is also explained at paragraph 87 of Applicant's patent application.

Applicant has claimed using records of mixed structured and unstructured data, using linguistic information to extract multi-dimensional relational fact from the unstructured data (free text), and then relating those relational facts to the structured data. Nothing Haug or the combination of references assembles this series of steps.

In addition to Haug, the Examiner has rejected claims based on combinations of Haug, Chen and Smith. The prior art from Chen is distinguishable in the same way as Haug. Chen teaches structuring unstructured data in this way: "call agent notations get converted into one or more "core concepts."" These core concepts are similarly one dimensional. Chen talks about the possibility of extracting more than one "core concept" but this still falls short. Even if you extract two core concepts, there is no necessary relationship between them other than co-occurrence in a record. For example, if the concepts extracted were "police officer" and "shooting," it would be impossible to determine whether the police officer was shot by someone, or whether the officer shot someone else himself, or some other occurrence happened. In contrast, the invention's extracted relational facts provide multiple roles and relationships that make clear who did what, to whom, when, etc.

Smith teaches the integration of extractions from multiple disparate unstructured text data

Serial No. 10/729,417  
Wakefield, et al.  
Filed December 5, 2003

**Response to Office Action**

Examiner: Phuong Thao Cao  
Group Art Unit: 2164

sources. Additionally, the type of information extraction it discusses is neither the concept extraction of Haug and Chen, but is instead entity extraction. Entity extraction is also one dimensional like concept extraction. It simply tells you that a person or thing exists in a piece of text. It tells you nothing about what he or it did, when, where, why, etc. It is very different from relational fact extraction, and again that difference provides key performance differences in the invention.

For these reasons, withdrawal of each of the rejections is appropriate. Reconsideration is requested.

Respectfully submitted this 1<sup>st</sup> day of December, 2006.



Daniel P. McCarthy  
Reg. No. 36,600  
P.O. Box 71550  
Salt Lake City, UT 84171-0550  
Tel 801 661 8998  
Fax 435 615 9669  
Email: dmccart@xmission.com